Anomalous Activity in the Bitcoin Blockchain

Andrew Marble

Team rbitr

andrewmarble@gmail.com

Intro – Problem

- Aim: To use Machine Learning to study patterns publicly visible in the bitcoin blockchain to identify indicators of **large-scale fraud, cyber ransomware attacks or other indicators of use for public safety**, safety response or other public policy implementations.
- Background: All bitcoin transactions are stored anonymously in the bitcoin blockchain. Individual addresses (payment endpoints) and transactions can be explored using existing tools, e.g. blockchain.info. Machine learning approaches look for patterns in the blockchain as a whole - features of blocks, transactions, or address payment histories that may be markers of criminal activity.

Intro – About Me

- Electrical Engineer
- Background in digital signal processing and time series analysis
- Interest in blockchain and large data set analysis
- Live in Ottawa

andrewmarble@gmail.com

Andrew Marble

Blockchain - Bitcoin

Blocks: records of transactions. 519481 as of Sunday evening, a new one mined every 10 min

Transactions: moving money between addresses. 300+ Million

Addresses: unique identifiers where bitcoin is stored – Hundreds of millions

Overall blockchain record is currently 140+ GB

Concept

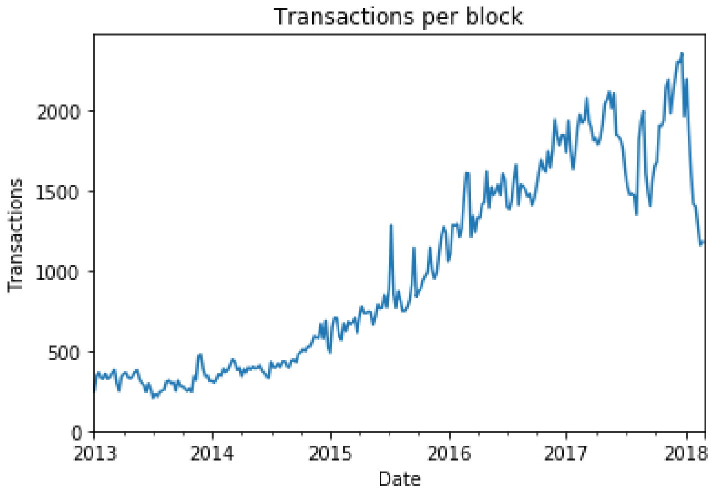
- Tableau for blockchain: visually summarize a large data set, and provide zoom / drill down capability allowing suspect or unusual transactions to be examined
- We want to begin by examining the properties of blocks and transactions to see what is normal and what is unusual

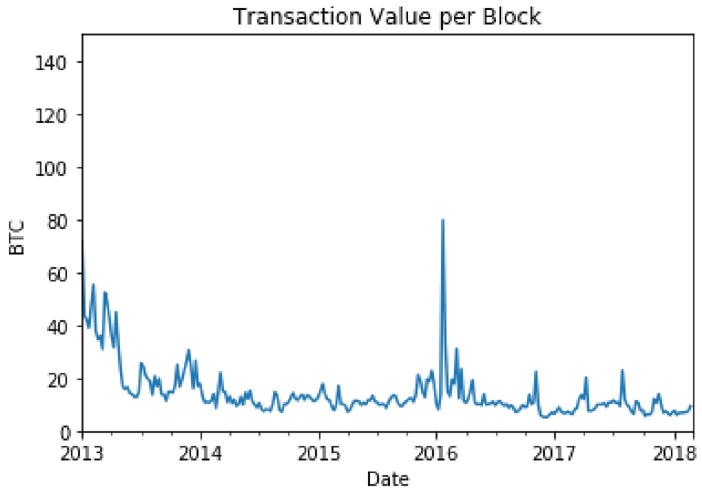
Tool Development

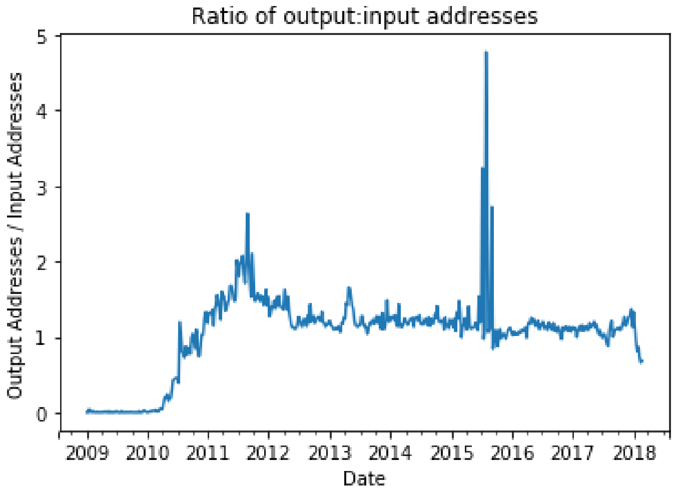
- Amazon Web Services “Deep Learning” machine image - 2x4 Xeon E5-2686 CPU, 61 GB RAM
- Build of Block-Sci blockchain analysis tool
- Full Bitcoin Node
- Prototype / trial implemented in python as a Jupyter notebook that combines code with results: github.com/rbitr/gcbc

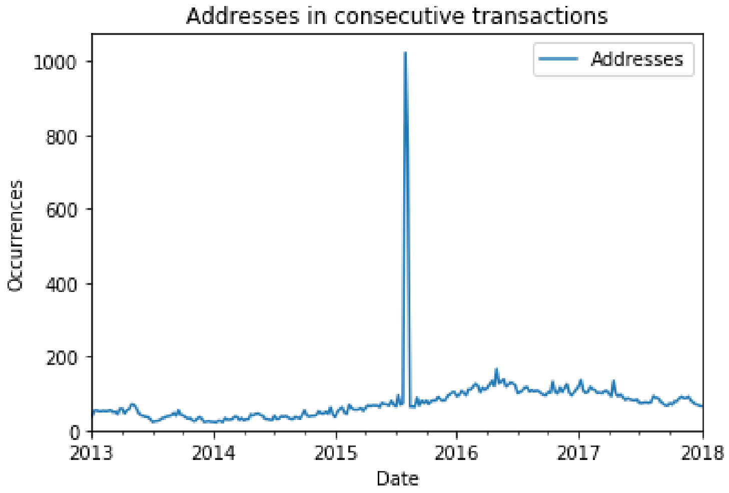
Block Features

Each represents a property of a block. Changes from the norm, in one or more features, or certain combinations of features may indicate activity warranting investigation


Transactions per block

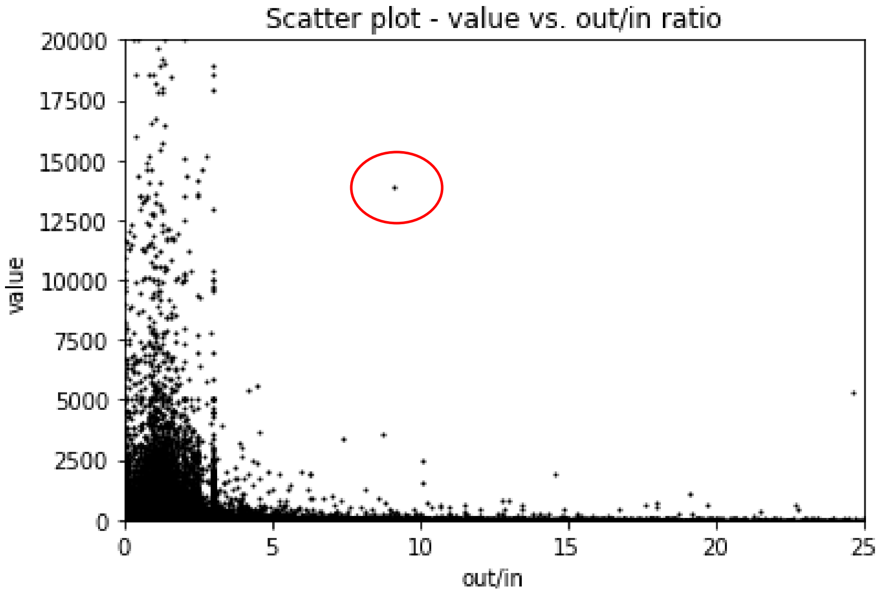

Transaction Value per Block

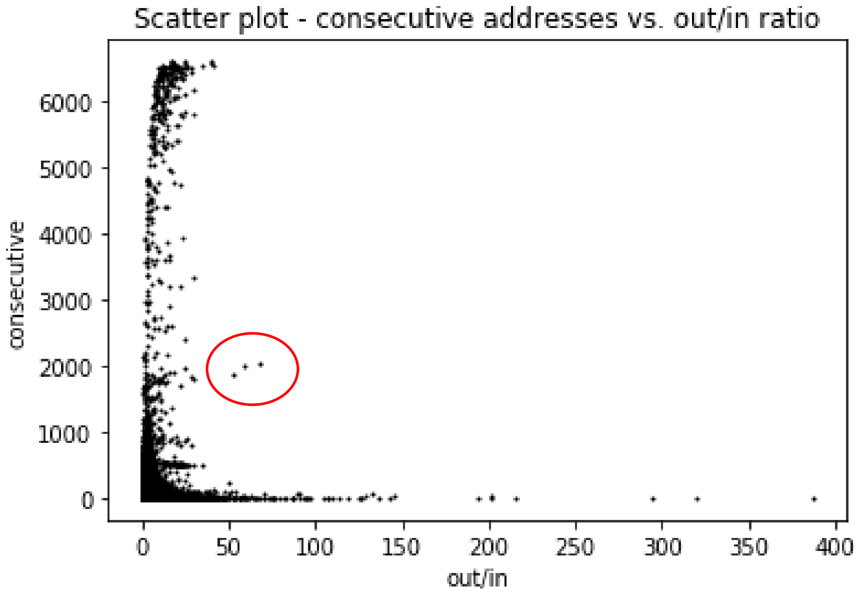

Ratio of output:input addresses


Addresses in consecutive transactions

Block Features - Outliers

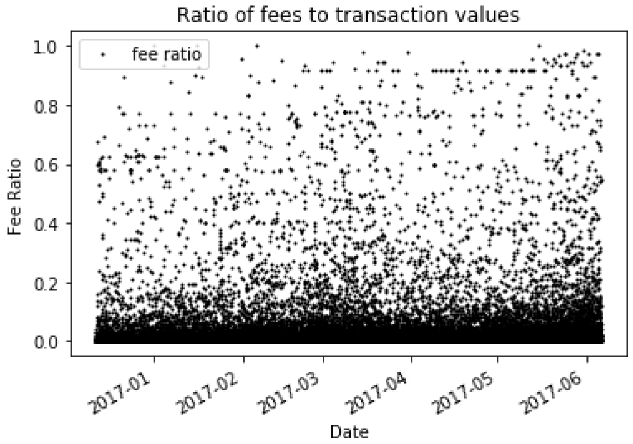
We can plot features against each other to look for combinations outside the norm. Each point represents a single block. Outliers – shown in red – sit away from the group and can be investigated

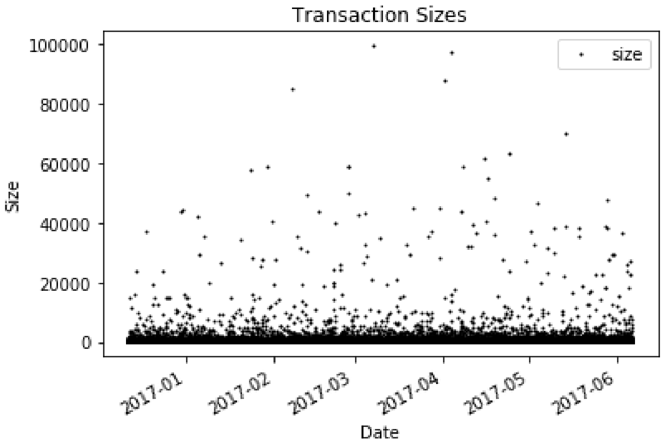


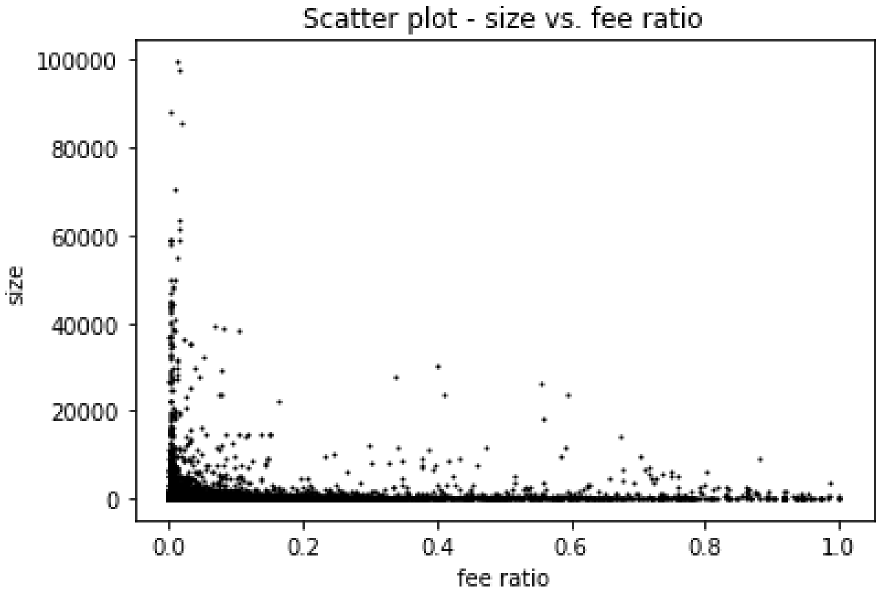


Transaction Features

Each represents a property of a transaction. As with blocks, outliers may indicate activity warranting investigation

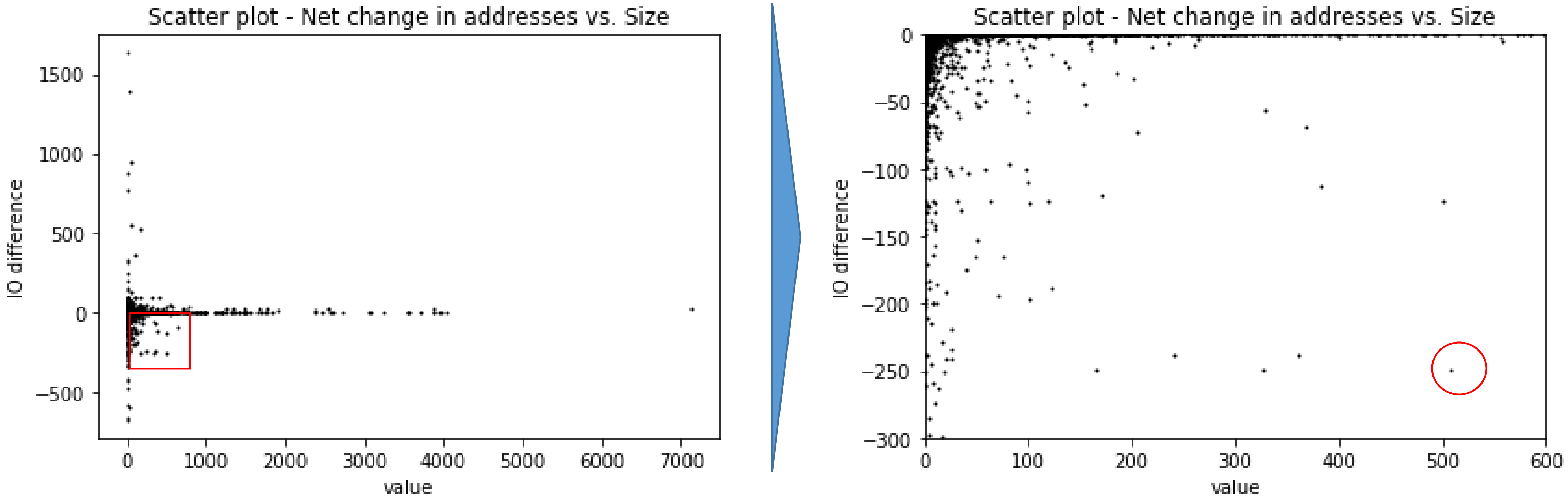






Transaction Features - Outliers

Plots of features can be used to zoom in on suspect transactions



Transaction e7f038f3ab7d061eaa1e6db4e89c840b3a984a8e15b509e432a83aaa21ecef70 is flagged. On Dec 12, 2016, a transaction consolidated 508 BTC from 250 addresses into a single address

Next Steps

- Automate some outlier detection – implementation is pre configured with python machine learning tools
- Build on feature set to include adjacent transactions, etc
- Training and validation with actual events or transactions

Conclusion

- Presented a method of flagging anomalous transactions or blocks that may signal public safety or fraudulent events
- Configured a hardware / software platform that can perform this analysis
- Demonstrated a proof of concept for a ‘Tableau for Blockchain’ analysis platform that can be used to flag suspicious activity